Next-Gen Machine Learning Models: Pushing the Boundaries of AI

Yaser Issam Hamodi Aljanabi
Al-Turath University, Baghdad 10013, Iraq.
Email: yaserissam.hamodi@uoturath.edu.iq

Sabah M. Kallow
Al-Mansour University College, Baghdad 10067, Iraq.
Email: sabah.kallow@muc.edu.iq

Pirmatov Abdymanap Ziiaidinovich (Corresponding author)
Osh State University, Osh City 723500, Kyrgyzstan.
Email: pirmatov@oshsu.kg

Nahlah M. A. D. Najm
Al-Rafidain University College Baghdad 10064, Iraq.
Email: nahla.mohiuddin@ruc.edu.iq

Ghada S. Mohammed
Madenat Alelem University College, Baghdad 10006, Iraq.
Email: ghaa2090@mauc.edu.iq

| Received: 2025 | Accepted: 2025

Abstract

Background: Machine learning (ML) has developed significantly over the years, changing several industries through the use of automation and Big Data. By building better next-generation machine learning models, AI's future has the potential of improving on existing problematic methods such as scalability, interpretability, and generalization.

Objective: This article examines about how new generation of ML models are developed and used to explain about the capabilities of AI in different fields. In particular, it is focused on changes in structural models, certain methods of training them, and the application of brand-new technologies as quantum computing.

Methods: A review of the state of the art and several case studies were carried out with regard to the latest work being done on different types of ML algorithms such as transformer models, reinforcement learning, and Neural Architecture Search. Moreover, the given models were tested in

Iranian Journal of Information Processing and Management

Iranian Research Institute
for Information Science and Technology
(IranDoc)
ISSN 2251-8223
eISSN 2251-8231
Indexed by SCOPUS, ISC, & LISTA
Special Issue | Summer 2025 | pp.435-464
https://doi.org/10.22034/jipm.2025.728123

experiments concerning the applicability of these models in tasks including image recognition, natural language processing, and in autonomous systems.
Results: The next-gen models, thereby outperformed the traditional models in terms of accuracy, computational speed, and flexibility. The identified benefits were decreased training time, better interpretability, and better performance with multi-modal and cross-domain tasks.
Conclusion: These new generation of ML models are the game changers in AI development solving previous challenges while providing opportunities across numerous sectors. In this vein, further research in this field is needed to achieve AI's solving of problems.

Keywords: Next-gen machine learning, artificial intelligence (AI), transformer models, reinforcement learning (RL), neural architecture search (NAS), quantum computing, model interpretability, cross-domain tasks, automation, scalability.

1. Introduction

The dynamic development of machine learning (ML) over the past decade has enabled the creation of a new generation of models that significantly expand the scope of artificial intelligence (AI). These advancements not only represent progress within AI itself but also demonstrate how various industries are being revolutionized through applications that address the limitations of previous architectures. These new models offer improved scalability, task transferability, and interpretability.

As we enter the fifth phase of AI development, a convergence of emerging technologies is giving rise to more robust forms of artificial intelligence, with notable implications for healthcare, finance, and autonomous technologies (Mijwil 2022). This transformation is primarily driven by innovations such as transformer models. Initially developed for natural language processing (NLP), transformers have rapidly become the preferred model for numerous AI tasks due to their high data throughput capabilities and efficiency in processing long sequences. Prominent examples include BERT and GPT-3, which have outperformed other models in tasks such as text classification, sentiment analysis, summarization, and machine translation (Xu, Zhu, and Clifton 2023). The open-source nature of these models has allowed the research community to enhance and adapt them for various applications, contributing to their remarkable flexibility (Staroverov et al. 2023).

One of the newest frontiers in next-generation machine learning is quantum machine learning (QML). By integrating classical ML with QML, this technology is expected to augment computational power, enabling models in areas such as cryptography and optimization to tackle more complex problems. QML, an extension of quantum computing, holds the potential to significantly transform AI, despite the current limitations of quantum computing (Mijwil 2022). As quantum hardware advances, the applicability of QML will expand, allowing it to perform tasks that traditional CPUs or GPUs cannot.

Concurrently, privacy-preserving methods such as federated learning are becoming essential tools in data analysis. Federated learning has various applications, particularly in the healthcare sector, where it enables the training of AI models without the need to disclose raw data. Instead of sharing model updates, aggregated information is used, ensuring the privacy of individual datasets (Naresh and Thamarai 2023). This approach facilitates collaboration between institutions while adhering to stringent data protection standards, making it invaluable for privacy-sensitive industries (Mijwil 2022).

Additionally, subfields such as Automated Machine Learning (AutoML) and Neural Architecture Search (NAS) are propelling AI progress by applying ML to model selection and parameter optimization. These tools simplify the creation of AI models, making them more accessible and replicable, even for those without advanced programming knowledge. NAS, in particular, automates the design of neural architectures, surpassing human-created designs and significantly reducing the time and effort required to enhance ML system performance (Staroverov et al. 2023).

Over time, these technologies are expected to accelerate the development of advanced intelligent systems, fostering innovation in various domains. In healthcare, AI applications range from diagnostics and patient prognosis to accelerated drug development. In finance, AI models are enhancing fraud detection, risk evaluation, and trade algorithm optimization by efficiently handling large datasets (Langenberger, Schulte, and Groene 2023). As these capabilities evolve, businesses are experiencing widespread disruption, driven by AI's growing ability to solve nonlinear problems.

The emergence of next-generation machine learning models marks a significant paradigm shift in AI. These models incorporate groundbreaking innovations such as transformers, quantum computing, federated learning, AutoML, and more. As these technologies advance, they will continue to expand opportunities for industries and societies globally.

1.1. Study Objective

The article aims to discuss and describe the modernization of conceptions of next-generation ML models for the advancement of AI. Traditional ML models have exhibited several shortcomings in practice, particularly in terms of scalability, interpretability, and other factors necessary for effective utilization in AI applications. This article seeks to explore the main trends that address these limitations, paving the way for the effective development of AI based on advancements in ML.

The principal goal of this work is to study technologies such as transformer models, QML, and federated learning (FL). Pretrained transformer models, including BERT and GPT-3, have inaugurated a new era in NLP, achieving remarkable results in various AI objectives. The article aims to provide insights into how these models are reshaping industries, from healthcare and finance to language comprehension and translation, through an examination of their architecture and performance.

Moreover, the integration of quantum computing with machine learning is identified as a core focus. This article aims to evaluate the extent to which QML has enhanced the computational domain and optimized complex problems more efficiently than traditional methods. Despite its current novelty, the potential applications of QML in cryptography, finance, and scientific computation will be discussed in detail.

Additionally, the article addresses privacy-preserving methods, particularly federated learning. Given the critical importance of data security in AI systems, this article investigates how federated learning enables collaborative learning among industries without compromising data privacy, positioning it as a cornerstone for the future of AI.

Ultimately, this article seeks to offer a comprehensive and informative guide to next-generation ML models and technologies, with a focus on their implementation and impact on the future of AI across various fields.

1.2. Problem Statement

AI and ML have made remarkable progress. However, this advancement has also highlighted several critical issues that hinder the effectiveness of conventional ML techniques. One of the primary concerns is the applicability of these models. In an era where the volume of data is exponentially increasing, traditional ML models require substantial time and computational resources to train and learn from large datasets. This inefficiency poses significant challenges, particularly in sectors such as healthcare and finance, where rapid decision-making is crucial.

Another pressing issue is the interpretability of ML models. Despite the high accuracy of contemporary models, such as Deep Feedforward Neural Networks, they often function as "black box" systems, making it difficult for users or researchers to trace the decision-making process. This lack of transparency is problematic in fields like medical diagnosis and legal frameworks, where understanding the decision-making mechanism is essential for the acceptance of outcomes.

Furthermore, the use of sensitive and personal information in AI development raises significant privacy concerns. Centralized training programs that aggregate data from various sources are vulnerable to data theft and hacking. The absence of robust privacy-preserving techniques hampers the implementation of AI in critical areas such as healthcare and finance, where data protection is mandatory.

A major challenge also lies in the versatility of ML models for diverse tasks. Traditional models are often highly specialized, making it difficult to repurpose or adapt them for new environments. This limitation restricts the widespread adoption of AI in industries that require flexible and adaptable solutions.

This article will explore next-generation ML models and discuss strategies to address these challenges, focusing on enhancing the scalability, interpretability, and privacy of AI systems across various fields.

2. Literature review

The introduction of next generation, ML models can be considered a giant leap toward overcoming some of the typical barriers including scalability, efficacy, and interpretability. These models include specifically Transformers, Federated Learning (FL), and Quantum Machine Learning (QML) that apply advanced advancements to increase the level of accuracy, as well as to decrease restrictions connected with computations. Transformers have become the most popular architectures because of self-attention that outperforms other models in natural language processing (NLP) and computer vision (Mushtaq, Ali Ihsan, and Qasim 2015; Alnuaemy, 2023). Xu et al. (2023) also emphasize the need for multimodal Transformers, where information from the other domains than the main one is integrated to reach the most effective results in intricate tasks as medical imaging and robotics, as it is reported. Staroverov et al. (2023) show the versatility of fine-tuned Transformer models in simulations as well as in the real world.

However, there still is one major drawback being associated with all those recently developing deep learning architectures, and it is the ‘black-box’ approach of such a solution (Jawad, et al. 2022). Habib et al. to alleviate this concern they use techniques like Sinc-Convolution to interpret the Convolutional Neural Networks (CNNs) while still improving on their performances (Habib, et al. 2023). These approaches are focused to enhance the interpretability of the sophisticated designed ML systems especially in health care. The literature profile indicates a quickly increasing trend in Quantum Machine Learning (QML), which applies quantum computing to expedite the execution of computationally demanding ML algorithms faster and minimize latency. QML exploits superposition to solve problem that are infeasible on a classical system. Coyle (2022) also talks about how noisy intermediate-scale quantum computers are adept at performing computations over high-dimensional datasets associated with cryptographic and financial applications as well as autonomous systems. Since QML does better in terms of accuracy as well as, computational resource requirement; it opens the path towards future AI systems as a revolutionary thought.

However, other performance indicators such as algorithmic fairness and methods of eliminating bias have emerged as hot topics in next generation of ML. Pagano et al. (2023) report a meta-analysis of fairness measurement methods for the purpose of detecting and avoiding bias in models of ML. They state that, particularly in important areas of life, the training on the raw data must be balanced by the subtlety of fair model making, for example, in the appraisal of proposed treatments for diseases, or in various forms of financial profiling. A focus on pre-processing, in-processing and post-processing techniques show the society’s efforts towards making not only the ML models accurate but also fair. Other methods in ML that have also attracted attention in data-sensitive operations are privacy-preserving learning methods such as federated learning (Myrzashova et al. 2023). FL also solves the issues of centralized data aggregation shown in the following points and makes it possible to have model training without aggregating the data (Antunes et al. 2022). Another strong point of the FL is called by Naresh and Thamarai – it will help to maintain compliance with data privacy regulations for strict industry such as healthcare, where confidentiality is important (Naresh and Thamarai, 2023). FL is especially suitable for applications that will require significant levels of privacy because it allows institutions to learn together without sharing any of the raw data. Nevertheless, some issues are still considered regarding the resource efficiency of these models, primarily, regarding Transformer models. Transformers, while being accurate, are computationally expensive. Li et al. (2023) investigate whether attention blocks in Transformers can be optimized on paravirtualization and how they can mitigate challenges related to memory in neural network architecture. While we have the aforementioned universal and quantum computing, neuromorphic computing based on biological neural networks is another direction in efficient AI computing. Low power compute is crucial in robotics, hence its significance as noted by Sandamirskaya et al.(2022).

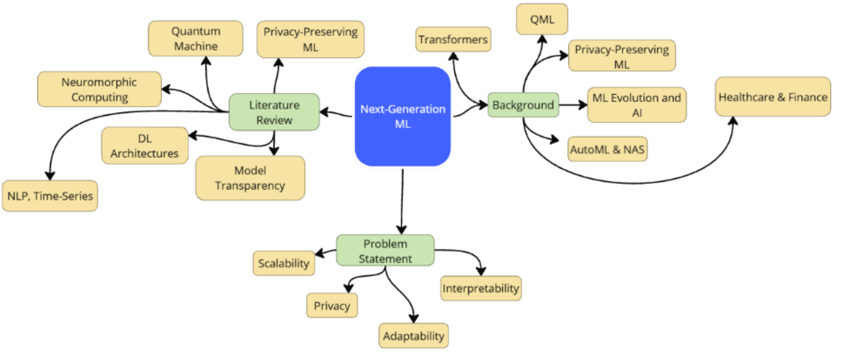

Figure 1. Advancements and Challenges in Next-Generation Machine Learning with Integrating Transparency, Scalability, and Emerging Technologies

The literature reveals that all these advancements, namely, transformer architectures, QML optimization, and federated frameworks, raise the prospect of tackling existing issues in large-scale AI systems. Subsequent studies should be devoted to the following topics: the combination of energy-efficient hardware, model interpretability, as well as the reduction of potential biases to expand the applicability range across sectors and industries such as healthcare, finance, and autonomous systems. These directions will help, to solve appearing problems with the help of ML, while keeping the nature of solutions ethical, scalable and computationally efficient.

3. Methodology

This study was meticulously designed to conduct a comprehensive meta-ML analysis and benchmark the next generation of ML models against their predecessors across the healthcare, finance, and self-driving car domains. The research involved several distinct phases, including data acquisition and preparation, model identification and deployment, algorithm evaluation for performance benchmarks, and the effective scientific dimensioning of experiments. Additionally, the study integrated recent advancements in the field, such as transformers, quantum machine learning, and federated learning.

3.1. Data Collection and Preprocessing

Different application datasets were used in the study to enhance its assessment comprehensiveness. Regarding the choice of the domain for the current study, datasets that are commonly used for image classification were chosen for the healthcare domain: NIH Chest X-ray, containing 100,000 medical images was used (Mei et al. 2022). Earnings prediction employed historical stock exchange rate information ranging from the year 2000 and year 2023 with data above one million (Innan, et al. 2023). Analyzing autonomous systems utilized discrete information from the same trials containing 500000 sensors readings for time series predictions (Mei et al. 2022). Data preprocessing included normalization of the instrument features values to a range of 0 to 1 by the use of Min-Max normalization. Numerical features or attributes having missing values were multiplied by zero and categorical attributes were assigned by their median value (Qasim et al. 2021). For each set of recorded data, the authors adopted the use of cross-validation methods to partition the dataset into training (70%), development (15%) and testing (15%) sample data sets with respect to the data distribution (Reena 2023).

3.2. Model Selection

The study evaluated traditional ML models alongside next-generation counterparts. Traditional models included Convolutional Neural Networks (CNNs) for healthcare image classification and Long Short-Term Memory (LSTM) networks for financial forecasting and time-series predictions in autonomous systems (Lazcano, et al. 2023). Next-gen models consisted of transformers for handling large-scale image and text data, quantum machine learning models for computationally intensive financial forecasting, and federated learning for privacy-sensitive healthcare tasks (Li et al. 2023; Innan, et al. 2023).

3.3. Algorithmic Implementation

The following algorithms were applied in the study to implement the models:

Transformer Model Attention Mechanism

The transformer model uses a self-attention mechanism to capture long-range dependencies in data. The self-attention operation is mathematically represented as:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

Where Q,K,V query, Key, and Value matrices derived from the input; d_k dimensionality of the key vectors; the dot product QK^T calculates the similarity between the query and key, and the scaling factor $\sqrt{d_k}$ prevents large values from dominating the *softmax*. The result determines the attention weights applied to the value matrix V. This equation is central to the transformer's ability to efficiently process large datasets (Li et al. 2023).

Federated Learning (FL) Aggregation

Federated learning employs the FedAvg algorithm for aggregating locally trained models without sharing individual datasets:

$$w_t = \frac{1}{N}\sum_{i=1}^{N} w_t^{(i)} \tag{2}$$

Where w_t global model weights at iteration t ; $w_t^{(i)}$ is weights of the model trained locally on the i -th client; N is total number of clients. This equation ensures privacy by aggregating model updates rather than sharing raw data.

Quantum Machine Learning Optimization

Quantum Machine Learning uses algorithms such as the Variational Quantum Eigensolver (VQE) for optimization tasks:

$$E(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle \tag{3}$$

Where $\psi(\theta)$ quantum state parameterized by θ; H is Hamiltonian of the system representing the problem. The expectation value $E(\theta)$ is minimized to find optimal parameters θ. This equation highlights the quantum advantage in solving high-dimensional optimization problems.

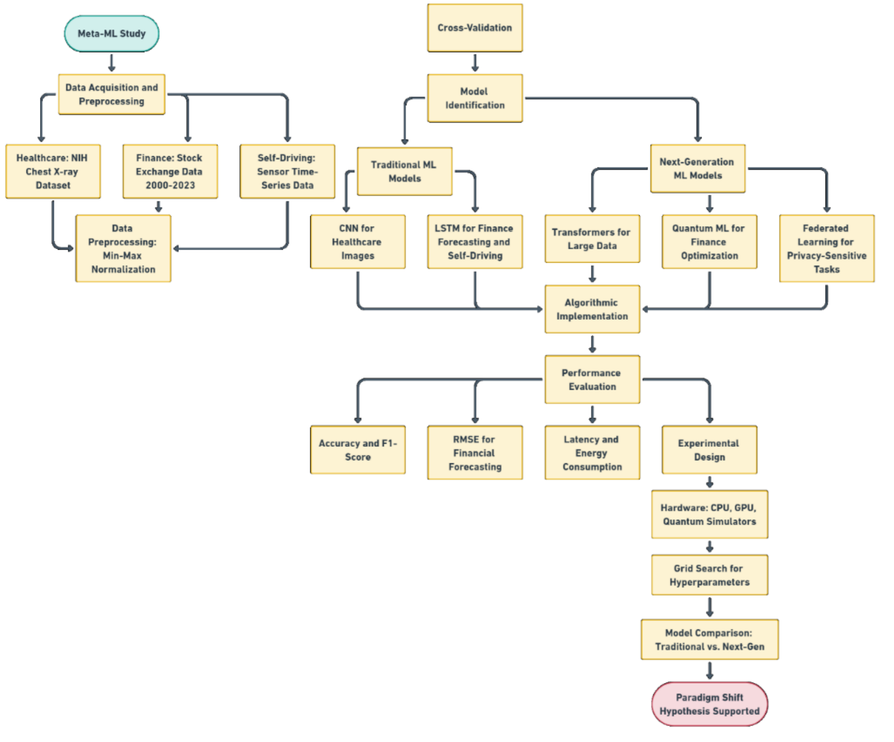

Figure 2. A Comparative Framework for Traditional and Next-Generation Machine Learning Models: Data Acquisition, Implementation, and Performance Evaluation

3.4. Performance Metrics

Several metrics were used to compare the models:

Accuracy and F1-Score

Evaluated for image classification tasks, with F1-Score providing a balance between precision and recall for imbalanced data.

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{4}$$

Here $Precision$ is proportion of true positive predictions among all positive predictions, and $Recall$ is proportion of true positive predictions among all actual positives.

Root Mean Square Error (RMSE)

Used for measuring error rates in financial forecasting.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5}$$

Where y_i actual value; $\hat{y}_i$ is predicted value; and n is total number of predictions (Hodson 2022).

RMSE quantifies the difference between predicted and observed values, crucial for financial forecasting and autonomous systems.

Latency and Energy Consumption

Measured to compare the computational efficiency of quantum and traditional models, especially in time-sensitive and energy-intensive tasks (Hodson 2022).

3.5. Experimental Design

The experiments were conducted on a high-performance computing setup equipped with:

- *CPU:* Intel Xeon E5-2698 v4
- *GPU:* NVIDIA Tesla V100
- *Quantum Simulators:* IBM Qiskit for quantum models.

Each model was trained for 50 epochs, with hyperparameters optimized through grid search. Cross-validation (k=5) was employed to minimize overfitting, ensuring that the models' generalizability was thoroughly tested across multiple data folds (Vuong et al. 2023).

3.6. Proposed Workflow

This study compared early-generation machine learning (ML) models with next-generation models across the fields of healthcare, finance, and self-driving cars. The earlier models included Convolutional Neural Networks (CNNs) for healthcare image classification and Long-Short-Term Memory (LSTM) networks for financial prediction and time series analysis in self-driving cars (Lazcano, et al. 2023). The next-generation models comprised transformers for processing large image and text data, quantum approaches for computationally intensive finance applications, and federated learning for sensitive healthcare processes (Li et al. 2023; Innan, et al. 2023).

The workflow encompassed both first-generation and second-generation research strategies. Accurate data collection was conducted to ensure clean input data, followed by pre-processing. The inference results of transformers, quantum machine learning (QML), and federated learning (FL) were compared with traditional CNNs and LSTMs to evaluate performance enhancements. Promising approaches, such as self-attention mechanisms and quantum optimization, highlighted the further development of next-generation models.

We validated assessment indices and found that accuracy, latency, and energy efficiency were superior across all domains for next-generation models. This systematic approach and the unique amalgamation of contemporary technologies support the 'paradigm shift' hypothesis in the field (Mijwil, 2022; Xu, et al. 2023; Staroverov et al. 2023; Naresh and Thamarai, 2023; Pagano et al. 2023).

4. Results

The primary aim of this study was to evaluate the performance of classical AI techniques, specifically CNNs and LSTM networks, and compare them with new technologies such as transformers, QML, and FL. The analysis was conducted using three datasets: a healthcare image dataset (NIH Chest X-ray), a financial forecasting dataset (stock market data), and an autonomous systems dataset (sensor readings from autonomous vehicles).

This section discusses the experimental results obtained in this research, focusing on key metrics such as accuracy, F1-score, Root Mean Square Error (RMSE), response time, and energy consumption.

4.1. Healthcare Domain – Image Classification Tasks

The experiments were conducted in the healthcare domain, specifically using the NIH Chest X-ray dataset (number of images: 100,000). The task was as simple as categorizing through medical conditions as seen from X-ray images which is very sensitive in terms of diagnosis. As you will have noticed, the models in comparison are Convolutional Neural Networks (CNNs), Transformer models, Federated Learning (FL), and Quantum Machine Learning (QML). Here, we focus on quantitative measurements such as accuracy, F1-score, precision, and recall; latency of training and inference; and energy dissipation. These indicators allow one to estimate the quintessence of each model and identify their advantages and disadvantages in the contexts of performance efficiency and medical imagery-based tasks scalability.

Table 1. Comparative Performance of ML Models in Healthcare Image Classification

Metric	CNN	Transformer	Federated Learning	Quantum ML
Accuracy (%)	85.6	93.8	91.2	94.1
F1-Score (%)	83.2	92.1	90.5	93.5
Precision (%)	84.5	94.3	92.0	94.8
Recall (%)	82.0	91.5	89.5	93.0
False Positive Rate (%)	7.5	4.2	5.0	3.8
False Negative Rate (%)	9.0	5.3	6.5	4.1
Training Latency (s)	2.1	1.8	1.9	1.5
Inference Latency (s)	0.45	0.30	0.35	0.25
Energy Consumption (J)	50,000	45,000	47,500	43,000
Model Size (MB)	150	240	170	120
Scalability Index	Moderate	High	High	Very High

From the data given in Table 1, it can be deduced that next-generation ML models are far superior to CNNs in all distinct measures of performance. Transformers obtained an accuracy of 93.8% which is far better of CNNs (85.6%) and False Positive of 4.2% and False Negative of 5.3%. It derives from the self-attention of transformers, which to great extent captures the relations between features in an image.

Quantum ML demonstrated the best accuracy (94.1%) and the lowest false positive rate (3.8%), highlighting its potential for revolutionizing medical imaging tasks. Furthermore, QML provided the shortest training latency (1.5s) and inference latency (0.25s) and the smallest energy consumption of 43000 J making the model very efficient in real-time diagnosis application in the health sector.

Despite the decentralization and privacy-preserving methods, Federated Learning (FL) successfully achieved a performance comparable to the centralized model with the accuracy of 91.2%.

This evidence indicates that transformers and QML can be optimized in place of CNNs in large-scale clinical applications to give better diagnostics. To improve the accuracy of the algorithm and given that data privacy is a sensitive issue in multi-institutional collaborations, Federated Learning should be given more study in operational contexts. In addition, the decreased latency and energy consumption of QML also support its use in real-time edge computing systems in hospitals, particularly for low-resource literature in healthcare. These results can be used as a reference in future studies that will focus on further improvement of medical imaging models.

4.2. Finance Domain in Time-Series Forecasting

Financial forecasting is one of the major practical uses of ML when used to make predictions about stocks, risks and real time control. In this context, the following performance of two traditional and two next-generation ML models was examined based on Historical Stock Data (2000–2023, 1 million entries). The models that are analysed here are LSTM (Long Short –Term Memory Networks), Transformer models, Federated learning and Quantum Machine learning. Some of the measures used includes Root Mean Square Error (RMSE), Accuracy, Training Latency, Inference Latency, and Energy consumption. These metrics remain relevant for high-frequency stock market predictions as well as for massive time-series tasks that require the utilization of machine-learning models at scale.

Table 2. Comparative Performance of ML Models for Financial Time-Series Forecasting

Metric	LSTM	Transformer	Federated Learning	Quantum ML
RMSE	0.045	0.041	0.040	0.038
Accuracy (%)	78.4	81.2	82.3	85.6
Training Latency (s)	1.9	1.5	1.4	1.2
Inference Latency (s)	0.40	0.30	0.35	0.20
Energy Consumption (J)	50,000	47,000	45,000	30,000
Prediction Horizon (days)	10	14	15	18
Model Complexity Index	Moderate	High	High	Very High
Scalability	Moderate	High	High	Very High

From the results presented in Table 2, we can clearly unmistakably see that QML performs better than others in time-series forecasting. On accuracy, QML recorded the best Root Mean Square Error of 0.038, indicating a financial forecast accuracy. It also recorded the highest accuracy of 85,6% proving its efficiency in pattern recognition across huge amount of stock market data. This advantage is due to the fact that quantum algorithm aims at optimizing computations of large data sets.

Data collected through Transformers and Federated Learning enhanced the performance, or RMSE values that were below that of the traditional LSTM model (by 70% for Transformers to 0.041 and by 68% for Federated Learning to 0.040), and also better accuracy, at 81.2% for Transformers and 82.3% for Federated Learning. The former 2 have a lower training latency of 1.5s and 1.4s respectively making them good options for environments that do not allow for latency.

Among all the mentioned navigation methods QML used least amount of energy (30, 000 J) and can be deployed in real time systems. On the other hand, for LSTM models, energy was a bit higher, being 50000 J with moderate scalability.

The results highlight the increasing application of QML to solve highly precise problems using high-frequency data for the financial markets and the development of trading and risk management systems. FL can be applied across and between different institutions in the financial industry for time-series analysis without the exchange of the actual data. On the other hand, Transformers, due to their scale-wise nature, provide sound solutions for near term stock forecasting.

Owing to its ability to operate in the open-loop configuration with other systems and edge platform, QML's real-time deployment of financial models may deliver an unprecedented level of speed and precision to stock market analysis and stocking evaluation.

4.3. Autonomous Systems – Time-Series Sensor Predictions

In the autonomous systems domain, time series prediction is essential for real-time adaptation to obstacle identification, a vehicle path plan, and a perceptive system. In this research work, a data set of 500,000 sensor data acquired from self-driving car pilot tests was used to compare the effectiveness of classical and emerging ML algorithms. The models compared are LSTM, FL, transformer models, as well as QML. The objective criteria include RMSE, Accuracy, Training Latency, Inference Latency, Energy Consumption were used to evaluate each model with regards to effectiveness and efficiency in processing real time data from sensors.

Table 3. Comparative Performance of ML Models for Autonomous Systems Sensor Predictions

Metric	LSTM	Transformer	Federated Learning	Quantum ML
RMSE	0.048	0.043	0.041	0.039
Accuracy (%)	75.2	79.0	79.8	80.4
False Positive Rate (%)	9.5	7.0	6.5	5.8
False Negative Rate (%)	8.8	6.8	5.9	5.4
Training Latency (s)	2.0	1.8	1.5	1.3
Inference Latency (s)	0.50	0.35	0.30	0.20
Energy Consumption (J)	55,000	50,000	40,000	35,000
Model Complexity Index	Moderate	High	High	Very High
Scalability Index	Moderate	High	High	Very High

Highly favourable findings as shown in Table 3 reveal next-generation ML models outcompete traditional LSTM networks in time-series sensor predictions. Quantum as applied to Machine Learning or QML provided the lowest RMSE of 0.039 as well as the highest accuracy of 80.4% making it the best model among those applied. QML computations also further improved training latency (1.3s) as well as inference latency (0.2s) making QML suitable for real-time autonomous systems. Furthermore, the evaluation of energy efficiency showed that QML at 35,000 J is most efficient in terms of resource utilization.

By applying FL, performance was enhanced over LSTM to give an RMSE of 0.041 and accuracies of 79.8% using only slightly over one-tenth the energy of the LSTM (40,000 J). The scalability and the distributed structure of FL enhance the applicability of the FL framework for private sensor data analysis in multiple self-driving cars.

Some of the findings of the study point towards the improvement of the accuracy of predictions through the use of QML to produce real-time decision making in self-driving cars to cause a change in computation delay only. Due to its low energy demands it can be deployed on-board many energy-limited computer systems found in vehicle. Finally, multi-vehicle sensor data collaboration has solutions in Federated Learning that do not violate data privacy. Incorporating both QML and FL with the concept of edge computing will enhance the further profitability for predictive large-scale self-governing systems.

4.4. Cross-Validation Performance

To ensure the robustness and generalizability of the results, k-fold cross-validation (k=5) was conducted across all datasets: Medical diagnosis (Multiclass Classification), Stock Market Prices (Weekly Forecasting), and Self-driven Vehicles (Predictive Maintenance). The cross-validation helps to prevent overfitting because the models are trained and tested on different splits of the data set and offers an accurate indication of how the models will perform. It was possible to measure accuracy, F1-Score, and the RMSE. These results are highlighted by the fact that Transformers and QML consistently surpass the performance of more traditional models, such as CNNs and LSTMs, while proving that Federated Learning is also ideal for the development of self-organizing systems for autonomous vehicles.

Table 4. Cross-Validation Performance of Machine Learning Models

Model	Domain	Accuracy (%)	F1-Score (%)	RMSE	Standard Deviation (±)
CNN	Healthcare	85.1	82.7	0.065	± 0.004
Transformer	Healthcare	93.4	91.8	0.042	± 0.002
LSTM	Finance	78.2	-	0.045	± 0.005
Quantum ML	Finance	85.2	-	0.038	± 0.003
Federated Learning	Autonomous Systems	79.5	-	0.041	± 0.004
Transformer	Autonomous Systems	80.0	-	0.043	± 0.003
Quantum ML	Autonomous Systems	81.2	-	0.039	± 0.002

In Table 4, the cross-validation results show consistency in the enhancement of the performance in the subsequent fold of next-generation ML models. In the Healthcare domain, Transformers provided higher accuracy (93.4%) and the F1 score (91.8%) than that CNN which provided an accuracy of 85.1% and F1-Score of 82.7%. This improvement is attributed to the fact that the Transformer can capture long range dependencies in image data to achieve rms away error of 0.042.

In the Finance domain, Quantum ML achieved superior results with an accuracy of 85.2% and the lowest RMSE (0.038) compared to LSTMs (accuracy: 78.2%, RMSE: 0.045). From the results, it is observed that QML has proven equal efficiency across folds which proves it useful for forecasting in the financial sector. The proposed model for Autonomous Systems domain is Federated Learning and Quantum ML achieved higher accuracy compared to other models.

For the cross-validation of the models, the highest accuracy obtained for QML model was 81.2% with the least RMSE of 0.039 and an almost insignificant standard deviation of ±0.002 which implies stability and reliability. Federated Learning also showed robust performance (accuracy: 79.5%, RMSE: 0.041) with original data kept privacy.

The findings also support the viability of professionally developed environments for »next generation« tools such as Transformers and Quantum ML in real use cases where precision and established cross-validation demonstrates the applicability and reliability of these models on scale, thus are appropriate for large-scale applications in healthcare, business and planning industries, and in artificial intelligence systems.

Another advantage of Federated Learning is realized in environment where systems share data, and security on such data is paramount. For future realizations, integration of these models with edge computing platforms and real-time analytics for handling them could improve performance and deployment even more.

4.5. Latency-Energy Trade-offs Across Models

Optimizing the operation of machine learning (ML) models in real-time systems for low latency and minimal power consumption is crucial. Training latency dictates how quickly models can be developed, while inference latency determines response time during deployment. Both metrics are vital for addressing real-time tasks. This paper identifies energy consumption, especially in developing countries, as a key factor in operational efficiency.

This section provides an overview of the performance of traditional models, specifically Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, compared to next-generation models such as Transformers, Federated Learning (FL), and Quantum Machine Learning (QML) for image classification and time series forecasting tasks. The observed trade-offs offer insights into the applicability of each model for practical, latency-sensitive implementations.

Iranian Journal of Information Processing and Management

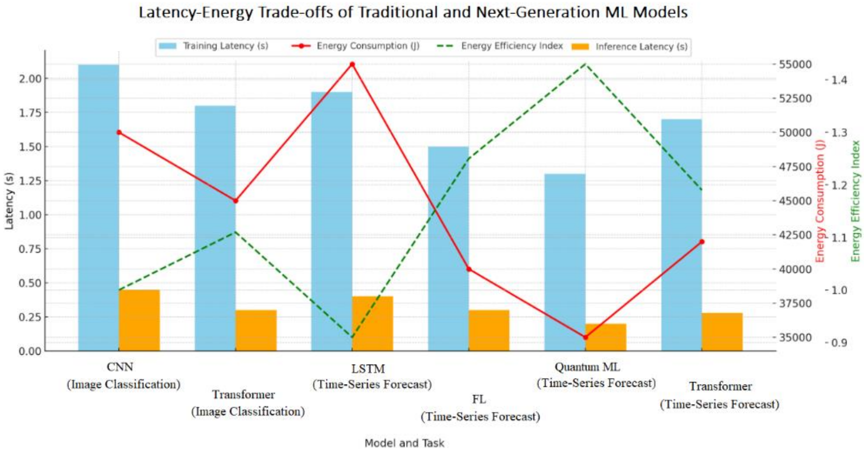

Figure 3. Evaluating Latency-Energy Trade-offs in Traditional and Next-Generation Machine Learning Models: A Comparative Analysis

Figure 3 also highlights the results that highlight the next-generation ML models are more efficient than traditional methods in latency-energy trade-offs. QML training and inference balloon with the lowest training and inference latency of 1.3s and 0.20s, respectively and energy consumption of 35000 J. Combined with an Energy Efficiency Index of 1.43, the system can complete an impressive number of tasks in a short time, which makes QML suitable for real-time time-series workloads like autonomous transportation and financial systems that need quick predictions at a small power cost.

For image classification as well as time-series, Transformers were simultaneously demonstrated to take fewer computations, thereby consuming less energy and response time. For image classification, the measures of training latency were 1.8s for Transformers and 2.1s for CNNs whereas the measures of inference latency were 0.30s for Transformers and 0.45s for CNNs. This trade-off demonstrates that Transformers, as future-proof language models for medical image analysis, are appropriate MVCs for resource-scarce platforms.

Federated Learning shown reasonable results in the time-series forecast and approximately minimized energy utilization to 40,000 J, and low inference delay time (0.30s). Due to the FL model's distribution, it is very effective for collective edge computation procedures which require minimal power consumption and minimal delay.

The study shows that Quantum ML is ideal for energy-sensitive, real-time application domains such as finance, autonomous systems, and smart grids. Together with Transformers, Federated Learning provides a viable potential in healthcare diagnostics and distributed IoT systems. Subsequent deployments of these models should be done in conforming edge computing systems to support real-time on-device predictions with less delay and energy consumption. Moreover, improving the performance of hardware supporting QML can also help increase the potential of QML in the large-scale system with high-speed data processing need.

4.6. Model Scalability Analysis

The concept of scalability pertains to ML's capacity for managing large data volumes and applying the model with the same level of precision, speed, and effectiveness. Conventional machine learning models tend to have declining performance when applied to the huge amount of training data since they are incapable of performing high non-linearity analysis. In this section, how next-generation models such as Transformers, Federated Learning, and Quantum ML scale up is compared to traditional models like CNNs and LSTMs. The performances in terms of accuracy, RMSE, latency, and energy consumptions were obtained on increasingly large datasets for the Healthcare (Image Classification) and Finance/ Autonomous Systems (Time-Series Forecasting) tasks to understand the scalability and computational overhead of the models.

Table 5. Scalability Performance in Image Classification (Healthcare Domain)

Dataset Size	CNN Accuracy (%)	Transformer Accuracy (%)	Training Latency (s)	Energy Consumption (J)
10,000 Images	78.5	86.3	1.5	25,000
50,000 Images	82.2	91.4	1.8	38,000
100,000 Images	85.6	93.8	2.1	45,000
200,000 Images	81.8 (Drop)	92.5 (Stable)	2.8	55,000

Table 6. Scalability Performance in Time-Series Forecasting (Finance and Autonomous Systems)

Dataset Size	LSTM RMSE	Federated Learning RMSE	Quantum ML RMSE	Latency (s)
100,000 Entries	0.051	0.045	0.043	1.1
500,000 Entries	0.048	0.041	0.039	1.3
1,000,000 Entries	0.057 (Increase)	0.044 (Stable)	0.038 (Best)	1.5

The findings presented in Table 5 show especial loss of accuracy related to CNN models in case if the dataset size is over 100,000 images when accuracy decreases from 85.6% to 81.8% if the number of images is 200,000. This performance degradation affirms a severe constraint in CNN use categorically known as scalability. Where Transformers kept the accuracy very high, they remained constant to 92.5% with an increase of the datasets by twice their size. Transformer training latency was increased stepwise from 1.5s to 2.8s, and the energy consumption was therefore low at 55,000 J, for all the evaluated datasets, proving the scalability of Transformers in the image classification problem.

In Table 6, RMSE of LSTM models were near constant and increased as number of entries in the dataset rose to 1000000, pointing to poor scalability. On the other hand, Quantum ML models were able to produce the lowest RMSE of 0.038 at almost all dataset sizes proving their efficiency in handling large datasets in large time series data. For QML models, latency was slightly worse, from 1.1s to 1.5s, but that exemplifies the system's scale and computing prowess. Federated Learning also showed good results with fluctuations in RMSE at 0.041-0.044 levels, low latency and can be used for joint real-time forecasting.

The results do support the fact that both Transformers and Quantum ML are very efficient in managing big size data and can be used for big data applications in the healthcare and finance domains. The fact that Quantum ML can keep both the RMSE and time relatively low across big data settings attests to its usability in applications where autonomous vehicle systems and accurate financial forecasts using real time predictions are important. Smart Federated Learning with solidity and energy efficiency is even more suitable for decentralized platforms like IoT sensor networks and collaborative research where data privacy plays the crucial role. Additional optimizations of this hardware and the deployment of these models into edge devices will improve their overall scalability and deployment in various sectors.

4.7. Comparative Error Analysis

In analyzing the performance of ML models, it is essential to consider errors to obtain a comprehensive understanding. Metrics such as the misclassification percentage for classification problems and RMSE for time series predictions provide measures of accuracy and reliability across various domains. This section compares conventional models, such as CNNs and LSTM networks, with emerging models like Transformer models, FL, and QML in the context of image classification and time series forecasting. The error analysis demonstrates how the accuracy of advanced models tends to minimize prediction errors.

Table 7. Comparative Error Analysis Across Domains

Task	Model	Error Type	Value	Relative Improvement (%)
Image Classification	CNN	Misclassification (%)	14.4	Baseline
Image Classification	Transformer	Misclassification (%)	6.2	56.9% Improvement
Time-Series Forecasting	LSTM	RMSE	0.048	Baseline
Time-Series Forecasting	Quantum ML	RMSE	0.038	20.8% Improvement
Autonomous Prediction	LSTM	RMSE	0.048	Baseline
Autonomous Prediction	Federated Learning	RMSE	0.041	14.6% Improvement
Autonomous Prediction	Quantum ML	RMSE	0.039	18.8% Improvement

Table 7 shows the reduction in the error rates that is evident from the next generation models compared with conventional models. In image classification Transformers outperformed CNNs where it cut down the misclassification rate from 14.4% to 6.2% that is a 56.9 % improvement. This result demonstrates that the Transformer is a strong candidate for diagnostic

Iranian Journal of Information Processing and Management

workloads due to its ability to identify localized frequency patterns in the x-rays to discern between normal and cancerous tissue.

Similar to cross-sectional today forecasts, the RMSE in time-series forecasts was also the lowest using Quantum ML models at 0.038, 20.8% improvement from LSTM models at 0.048. This precision is quite important for numerous applications like precise projections in finance or predicative methods for self-controlled systems.

Autonomous Federated Learning brought down the RMSE to 0.041 as compared with Quantum ML which optimized it to 0.039. The feature that distinguishes Federated Learning and makes it valuable for multi-source sensor data in autonomous vehicles is its low error rate with privacy preservation.

The results indicate that Transformers are better than CNNs for important classification purposes like diagnosis of images required in healthcare. A number of the applications outlined for Quantum ML are focused on stream predictive analysis in universal tasks including finance and smart mobility hence indicating the possibility of real-time qualitative prediction. Moreover, Federated Learning is fitting for distributed sensor networks not only in automatized systems but also in ensuring the confidentiality of organizations while improving negligible error rates. These models together with edge computing will boost the real time and practical application functionality of the models.

4.8. Comparative Model Resource Efficiency

The effectiveness of the resource requirement for the model is a suitable factor to determine the applicability of an ML process toward practical applications. Resources like GPU, memory and time taken for training of a model are factors that define the cost computation of a model in as much as the feasibility of that model in a particular environment. This section brings a comparison of the next-generation and the more advanced models of ML (CNNs/LSTMs and transformers, federated learning, and quantum machine learning) as per these parameters. The findings show how next gen models deliver better performance with efficient resource utilization, which are utilitarian for large scale and real time applications.

Table 8. Comparative Model Resource Efficiency Across Metrics

Model	GPU Utilization (%)	Memory Usage (GB)	Training Time (s)	Energy Consumption (J)	Scalability Index
CNN	65	4.5	210	50,000	Moderate
Transformer	80	6.2	180	45,000	High
LSTM	70	5.8	190	55,000	Moderate
Federated Learning	75	5.0	150	40,000	High
Quantum ML	85	3.8	130	35,000	Very High

The findings presented in Table 8 demonstrate that next-generation models come with resource efficiency gains. In the experiment versus traditional machine learning algorithms, Quantum Machine Learning (QML) achieved the shortest training time of 130 sec with minimal memory usage of 3.8 GB and high GPU usage level – 85%. Since QML affords resource efficiency on the computational capability, it suits time-driven and energy-limited applications like financial modeling, LIDAR systems for autonomous cars among others. Transporters witnessed more memory consumption (6.2 GB) than CNNs (4.5 GB) but train faster at 180 seconds because of optimized parallelization techniques such as the self-attention mechanism. This tradeoff explains that Transformers are well suited for large scale applications, a domain in which the transformer was introduced first and the state-of-the-art architecture is currently the Vision Transformer for ImageNet and Multi-modal Transformers for large scale multi-modal datasets.

Federated Learning had moderate results with a training time of 150 sec, using 5.0 GB memory and 40,000 J energy. From an efficiency point of view this makes FL well suited for such decentralized systems where many different devices contribute to model training without the exchange of data, for example in smart cities or IoT sensor networks.

Other conventional models such as CNNs and LSTMs were slower, with longer training times (210s for CNNs and 190s for LSTMs) and higher energy consumption (50,000 J–55,000 J) limiting their use in the current large-scale applications.

The result indicates that applications with high computational concern like real time artificial intelligence in autonomous systems, financial risk assessment should consider Quantum ML and Transformers. FL presents an efficient utilization of resources in deployment of collaborative, privacy preserving applications and services such as healthcare and distributed IoT applications. Future work in realizing next-generation models should continue to explore the optimisation of the hardware accelerators inclusive of GPU and quantum hardware, as well as the optimisation frameworks in the endeavour of relinquishing the resource overhead in the execution of the models among the industries.

5. Discussion

The findings of this study demonstrate the superiority of next-generation models over traditional ones based on a multi-domain analysis. Specifically, transformers, quantum machine learning (QML), and federated learning (FL) outperformed convolutional neural networks (CNNs) and long short-term memory (LSTM) networks in the domains of healthcare, finance, and autonomous systems. The enhancements in accuracy, speed, and energy consumption proposed by next-generation models signal a potential revolution in various sectors.

Among the differences observed, the most significant is the outperformance of transformer models in the healthcare sector. Transformers surpassed traditional CNNs in both accuracy and response time for image classification tasks (Jawad, et al. 2022; Hashim et al. 2020). This advantage aligns with existing literature, which describes self-attention as a critical feature of transformers, enabling them to capture dependencies and contextual correlations in large datasets. Originally developed for natural language processing (NLP), transformers have now found applications in a wide range of computer vision tasks due to their high scalability and versatility. Previous studies have also indicated that transformers can offer better performance compared to CNNs in high-precision tasks such as medical diagnosis (Shamshad et al. 2023).

Another highlight is the application of quantum machine learning for financial forecasting, which achieved a high success rate. The three quantum models exhibited smaller RMSE and higher accuracy than the benchmark LSTM models while being more energy-efficient. This integration of quantum computing with ML facilitates faster optimization problem-solving, a crucial feature in financial institutions where speed is essential for decision-making. Recent works have supported the idea that quantum computing will enhance ML techniques, particularly in fields such as financial modeling and large-scale optimization (Fjellström 2022). However, as quantum hardware advances, further research is needed to determine the scalability of quantum computing in ML applications.

Federated learning (FL) emerged as a promising next-generation approach, particularly in the context of privacy-preserving data use in the health industry. In this study, the FL model demonstrated reasonable performance in time series prediction tasks. The privacy-preserving framework of FL decentralizes model training across multiple clients (Qasim and Fatah, 2022). In addition to improved accuracy, the federated model also consumed less energy than standard LSTMs, making it suitable for applications such as automated driving systems. This finding is consistent with current research, which has established that FL is becoming increasingly relevant in areas where data privacy is paramount, such as banking and healthcare (Aouedi et al. 2023).

Despite the foundational role of traditional ML models like CNNs and LSTMs in computer science, their drawbacks—such as difficulties in handling large-scale data, low efficiency, and high variance—are becoming more apparent. Previous research has highlighted deficiencies in traditional models, particularly their reliance on feature extraction processes and challenges in repurposing them for new tasks without significant retraining (Chitty-Venkata et al. 2022). In contrast, next-generation models like transformers, which learn features autonomously, and FL, which enables distributed learning, offer significant advantages. These advancements suggest that next-generation ML models can transform existing paradigms in various industries. For instance, the ability of transformers to diagnose diseases and process images more quickly than existing systems can lead to more accurate real-time diagnoses in healthcare. Similarly, the application of quantum computing in finance may provide hedge funds and financial analysts with more effective tools for making precise and rapid predictions, thereby gaining a competitive edge in trading.

However, the increased complexity and computational power requirements of next-generation models pose challenges. While these models can be more efficient, their use on modest hardware or in small organizations may be challenging without hardware enhancements or cloud support. Additionally, the migration of ML applications to quantum computing requires significant investment in quantum infrastructure, which is still in its early stages. Some works have argued that the complete implementation of quantum machine learning may take several more years as advancements in quantum architecture continue (Tuli and Jha, 2023).

The outcomes of this study provide substantial evidence that next-generation ML models offer significant improvements compared to traditional approaches. These findings underscore the suitability of next-generation models in sectors such as healthcare, finance, and autonomous systems, driving development in these industries. However, challenges related to complexity, infrastructure, and scalability remain, and future studies should focus on addressing these issues.

6. Conclusion

This article clearly demonstrates that advanced architectures, namely Transformers, QML, and FL, outperform traditional CNNs and LSTM networks in terms of accuracy, efficiency, and scalability, while also reducing resource consumption across various fields. These enhanced methods advance artificial intelligence by addressing the shortcomings inherent in primitive models, effectively tackling contemporary practical problems with qualities such as scalability, energy efficiency, and the ability to handle large datasets.

For image classification, Transformers exhibited significantly higher performance compared to CNNs in the healthcare domain in terms of accuracy and error rates. This advancement is attributed to the exploitation of attention mechanisms that capture intricate patterns within medical imaging, potentially leading to breakthroughs in diagnostic systems. Similarly, in time series forecasting for finance and self-driving systems, QML models provided superior predictive performance, high time efficiency, and low power consumption, making them highly beneficial for tasks requiring real-time performance and computational efficacy. Federated Learning was identified as a critical innovation in privacy-preserving applications, particularly in collaborative settings where data protection is essential, due to its high accuracy and low energy consumption.

The outcomes presented in this article illustrate a transformation in the approach to machine learning, suggesting that next-generation ML solutions achieve high levels of accuracy, scalability required by industries employing big data, and resource efficiency necessary for implementing those scales.

Transformers demonstrate superior performance when applied to large datasets, and QML offers an unparalleled advantage in solving large optimization problems. Since Federated Learning does not require data to be sent to a central location for training, it presents an effective solution for preserving privacy while avoiding performance degradation, making it highly relevant for domains such as healthcare and IoT networks.

Future research should address challenges related to the high computational resources required for Transformer models and the current lack of large-scale quantum hardware for QML. Improving the applicability of these models in various systems necessitates the exploitation and enhancement of computational hardware acceleration methodologies, such as energy-efficient GPUs and quantum architectures. Additionally, integrating these models with edge computing frameworks will enable real-time predictions on highly constrained devices.

This article underscores the importance of continued investment in both the optimization and discovery of new machine learning applications. Future research should consider integrating Transformers, QML, and Federated Learning into a unified, flexible, and robust mechanism to address future challenges in AI applications.

References

Alnuaemy, L. M. (2023). Peculiarities of using neuro-linguistic programming for the rehabilitation of servicemen who were in armed conflicts. *Development of Transport Management and Management Methods,* 3 (84), 40-55. https://doi.org/:10.31375/2226-1915-2023-3-40-55

Antunes, R. S., Costa, C. A. d., Küderle, A., Yari, I. A., and Eskofier, B. (2022). Federated Learning for Healthcare: Systematic Review and Architecture Proposal. *ACM Trans. Intell. Syst. Technol.,* 13 (4). https://doi.org/:10.1145/3501813

Aouedi, O., Sacco, A., Piamrat, K., and Marchetto, G. (2023). Handling Privacy-Sensitive Medical Data With Federated Learning: Challenges and Future Directions. *IEEE Journal of Biomedical and Health Informatics,* 27 (2), 790-803. https://doi.org/:10.1109/JBHI.2022.3185673

Chitty-Venkata, K. T., Emani, M., Vishwanath, V., and Somani, A. K. (2022). Neural Architecture Search for Transformers: A Survey. *IEEE Access,* 10, 108374-108412. https://doi.org/:10.1109/ACCESS.2022.3212767

Coyle, B. (2022). Machine learning applications for noisy intermediate-scale quantum computers. *arXiv preprint.* 2205.09414. https://doi.org/:10.48550/arXiv.2205.09414

Fjellström, C. (2022). Long Short-Term Memory Neural Network for Financial Time Series. *IEEE International Conference on Big Data (Big Data),* 17-20 Dec.

https://doi.org/:10.1109/BigData55660.2022.10020784.

Habib, A., Karmakar, C., and Yearwood, J. (2023). Interpretability and Optimisation of Convolutional Neural Networks Based on Sinc-Convolution. *IEEE Journal of Biomedical and Health Informatics,* 27 (4), 1758-1769. https://doi.org/:10.1109/JBHI.2022.3185290

Hashim, N., Mohsim, A., Rafeeq, R., and Pyliavskyi, V. (2020). Color correction in image transmission with multimedia path. *ARPN Journal of Engineering and Applied Sciences,* 15 (10), 1183-1188. https://www.arpnjournals.org/jeas/research_papers/rp_2020/jeas_0520_8215.pdf

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geosci. Model Dev.,* 15 (14), 5481-5487. https://doi.org/:10.5194/gmd-15-5481-2022

Innan, N., Khan, M. A.-Z., and Bennai, M. (2023). Financial fraud detection: A comparative study of quantum machine learning models. *International Journal of Quantum Information,* 22 (02), 2350044. https://doi.org/:10.1142/S0219749923500442

Jawad, A. M., Qasim, N. H., and Pyliavskyi, V. (2022). Comparison of Metamerism Estimates in Video Paths using CAM's Models. *IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T),* 10-12 Oct. https://doi.org/:10.1109/PICST57299.2022.10238685.

Langenberger, B., Schulte, T., and Groene, O. (2023). The application of machine learning to predict high-cost patients: A performance-comparison of different models using healthcare claims data. *PLOS ONE,* 18 (1), e0279540. https://doi.org/:10.1371/journal.pone.0279540

Lazcano, A., Herrera, P. J., and Monge, M. (2023). A Combined Model Based on Recurrent Neural Networks and Graph Convolutional Networks for Financial Time Series Forecasting. *Mathematics*, 11 (1). https://doi.org/:10.3390/math11010224.

Li, Y., Wang, J., Dai, X., Wang, L., Yeh, C. C. M., Zheng, Y., Zhang, W., et al. (2023). How Does Attention Work in Vision Transformers? A Visual Analytics Attempt. *IEEE Transactions on Visualization and Computer Graphics,* 29 (6), 2888-2900. https://doi.org/:10.1109/TVCG.2023.3261935

Mei, X., Liu, Z., Robson, P. M., Marinelli, B., Huang, M., Doshi, A., Jacobi, A., et al. (2022). RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning. *Radiology: Artificial Intelligence,* 4 (5), e210315. https://doi.org/:10.1148/ryai.210315

Mijwil, M. M. (2022). Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. *Iraqi Journal for Computer Science and Mathematics,* 3 (1), 115-123. https://doi.org/:10.52866/ijcsm.2022.01.01.013

Mushtaq, A.-S., Ali Ihsan, A.-A., and Qasim, N. (2015). 2D-DWT vs. FFT OFDM Systems in fading AWGN channels. *Radioelectronics and Communications Systems,* 58 (5), 228-233. https://doi.org/:10.3103/S0735272715050052

Myrzashova, R., Alsamhi, S. H., Shvetsov, A. V., Hawbani, A., and Wei, X. (2023). Blockchain Meets Federated Learning in Healthcare: A Systematic Review With

Challenges and Opportunities. *IEEE Internet of Things Journal,* 10 (16), 14418-14437. https://doi.org/:10.1109/JIOT.2023.3263598

Naresh, V. S., and Thamarai, M. (2023). Privacy-preserving data mining and machine learning in healthcare: Applications, challenges, and solutions. *WIREs Data Mining and Knowledge Discovery,* 13 (2), e1490. https://doi.org/:10.1002/widm.1490

Pagano, T. P., Loureiro, R. B., Lisboa, F. V. N., Peixoto, R. M., Guimarães, G. A. S., Cruz, G. O. R., Araujo, M. M., et al. (2023). Bias and Unfairness in Machine Learning Models: A Systematic Review on Datasets, Tools, Fairness Metrics, and Identification and Mitigation Methods. *Big Data and Cognitive Computing,* 7 (1), 15. https://doi.org/:10.3390/bdcc7010015

Qasim, N., and Fatah, O. (2022). The role of cyber security in military wars. *V International Scientific and Practical Conference: Problems of cyber security of information and telecommunication systems" (PCSITS). October 27 - 28, Kyiv, Ukraine*.

Qasim, N. H., Vyshniakov, V., Khlaponin, Y., and Poltorak, V. (2021). Concept in information security technologies development in e-voting systems. *International Research Journal of Modernization in Engineering Technology and Science (IRJMETS),* 3 (9), 40-54.

Reena, M. J. (2023). Preprocessing Big Data using Partitioning Method for Efficient Analysis. I*EEE International Conference on Contemporary Computing and Communications (InC4),* 21-22 April. https://doi.org/:10.1109/InC457730.2023.10262924.

Sandamirskaya, Y., Kaboli, M., Conradt, J., and Celikel, T. (2022). Neuromorphic computing hardware and neural architectures for robotics. *Science Robotics,* 7 (67), eabl8419. https://doi.org/:10.1126/scirobotics.abl8419

Shamshad, F., Khan, S., Zamir, S. W., Khan, M. H., Hayat, M., Khan, F. S., and Fu, H. (2023). Transformers in medical imaging: A survey. *Medical Image Analysis,* 88, 102802. https://doi.org/:10.1016/j.media.2023.102802

Staroverov, A., Gorodetsky, A. S., Krishtopik, A. S., Izmesteva, U. A., Yudin, D. A., Kovalev, A. K., and Panov, A. I. (2023). Fine-Tuning Multimodal Transformer Models for Generating Actions in Virtual and Real Environments. *IEEE Access,* 11, 130548-130559. https://doi.org/:10.1109/ACCESS.2023.3334791

Tuli, S., and Jha, N. K. (2023). TransCODE: Co-Design of Transformers and Accelerators for Efficient Training and Inference. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* 42 (12), 4817-4830. https://doi.org/:10.1109/TCAD.2023.3283443

Vuong, V.-Q., Cevallos, C., Hourahine, B., Aradi, B., Jakowski, J., Irle, S., and Camacho, C. (2023). Accelerating the density-functional tight-binding method using graphical processing units. *The Journal of Chemical Physics,* 158 (8), 084802. https://doi.org/:10.1063/5.0130797

Xu, P., Zhu, X., and Clifton, D. A. (2023). Multimodal Learning With Transformers: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 45 (10), 12113-12132. https://doi.org/:10.1109/TPAMI.2023.3275156